Framing Effects and Risk Perception:

Testing Graphical Representations of Risk for the KIID

Remo Stössel[1], Anna Meier[2]

Department of Banking and Finance

University of Zurich

Abstract

In this paper we analyze which graphical representation of risk is most effective in supporting investors to assess the risk and return characteristics of a fund. Moreover, we test on which criteria the investors base their risk taking behavior. To this end we compare return bar charts and price line charts, combined with some additional information such as a risk scale or a gain and loss range.

We find that the risk communication with bar charts performs relatively well, except with regard to communicating the possibility of losses. Furthermore, we find that people generally underestimate risks and overestimate return. We additionally find that risk perception has the strongest influence on risk taking behavior, and in particular that a higher risk perception leads to less risk taking.

Key Words: risk communication, risk perception, risk comparability, risk awareness, loss aversion, performance path influence, risk taking behavior, graphical communication, KIID

JEL-classification: C93; G11

Acknowledgements: We are grateful to valuable comments by Thorsten Hens, Stefan Zeisberger and the participants of the seminar at the University of Zürich. Financial support from SNF-grant No CR12I1_149856 and from the UZH URPP Finreg is gratefully acknowledged.

[1] Department of Banking and Finance, University of Zurich, Plattenstrasse 32, 8032 Zürich, Switzerland
[2] Talbächliweg, 8048 Zürich, Switzerland

1 Introduction

The recent financial crisis has demonstrated that investors did not understand the risk of products they bought. A common mistake investors make is to take risk which they cannot psychologically handle and as a result they often sell in the worst possible moment. As for example Dalbar (2011) has estimated, the cost of those investment mistakes sums up to 4.3% per annum! The costs extend to wealth managers due to losing dissatisfied clients and in some cases to bear the costs of legal proceedings. To prevent such costs, in many financial markets the regulators are revising their rules for risk communication.

The European Undertakings for Collective Investments in Transferable Securities (UCITS) directive (European Parliament and European Council, 2009), for example, requires that for investment funds a key investor information document (KIID) needs to be prepared and be provided to all investors who consider investing in these funds. The KIID is standardized and contains the following components: objectives & investment policy, risk & reward profile, charges (fees), past performance and practical information. On the 1st of April 2014 the EU Parliament and Council agreed on a further regulation which extends the regulation for UCITS funds to a broader range of products, the so called packaged retail investment and insurance products (PRIIPs). Although the UCITS funds fall within the scope of the PRIIPs regulation they are exempt from the new key information document (KID) requirements for a transitional period of five years. A key difference between the KIID and the new PRIIP KID is that the former includes past performance while the later does not. Instead, the PRIIP KID includes forward looking scenario information. The difference shows that there is no consensus on what should be included in a key investment document, and how the information should be illustrated.

It is well known that different graphical formats can have a significant influence on the risk perception (Weber et al. (2005)), and are also selectively used by financial companies to better illustrate their financial performance (Beattie & Jones (2000)). For example the use of return bar charts result in framing effects, as a large gain (e.g. 100%) followed by a much smaller loss (e.g. 50%) may leave the impression that the over-all return is still positive, while it is in fact null.

In this paper we focus on the risk communication in the context of the KIID. To evaluate how different risk communication formats affect people, we did a large survey to compare the two presentation forms commonly used in practice, namely the return bar charts (used in the KIID) and price line charts. We evaluate them based on how *accurately* the risk and return characteristics can be assessed, and how well funds can be *compared* with them. Investors who can accurately assess the risks are less likely to take too much of it, and investors who can consistently assess risk and return can compare funds. These criteria are consistent with the proposal of the European Parliament and Council (2012), namely that investors should receive understandable and comparable disclosures. The risk and return characteristics the participants of our survey estimate are the expected return, volatility,

potential loss, and their perceived (past and future) risk. The additional information tested is a simple scale of risk to guide the investor's perception, the synthetic risk and reward indicator (SRRI) as currently used in the KIID, and a range of gains and losses to include in the graph. We furthermore test which criteria most strongly influenced the risk taking behavior, as finally the intention of the KIID is to nudge investors to take more appropriate investment decisions, i.e. not to take too much risk.

The results show that people perceive the risk to be highest with bar charts, that the SRRI supports comparability of funds, and that people severely under-estimate potential losses. We further show that the most important criteria for making an investment decision is the perceived risk, consistent with the findings of Weber et al. (2005). Based on these results we find that the communication currently used in the KIID performs relatively well, except with regard to losses. Price line charts best communicate the expected return and volatility, and the price line chart with the loss and gain range chart leads to the lowest underestimation of potential losses. Additional information on potential losses and gains would be required to support a more accurate awareness of the risk of losses, and likely also lead to a higher risk perception.

This paper is organized as follows. Section 2 and 3 provides a literature review and the research hypotheses respectively. Section 4 describes the survey design and data. The results are summarized in section 5. The discussion and conclusion are in sections 6 and 7 respectively.

2 Literature Review

In this section we briefly introduce the literature on risk communication from different fields, and the existing views on the different types of risk communication. Building on this introduction, the literature on the role of context and framing for risk communication is discussed. We continue with literature on the multi-dimensionality of risk perception, which needs to be taken into account when attempting to communicate risk. We conclude with a summary of the literature which shares the most similarities to this paper, namely Weber et al. (2005) and Diacon & Hasseldine (2007).

How to communicate risk is a challenge faced in many research areas, including the actuarial sciences (Duklan & Martin (2002)), medicine (Weber & Hilton (1990), Lipkus (2007), and Fagerlin et al. (2007)), psychology (Fischhoff (1995), Fischhoff et al. (1998)) and finance (Weber et al. (2005) and Veld & Veld-Merkoulova (2008)) among others. Some authors knowingly compare different presentations of probabilistic information used across different fields, such as Spiegelhalter et al. (2011). What they all have in common is that they try to find a presentation format which maximizes people's understanding of risk. While most presentation formats fall into one of the three categories graphical, verbal, and numerical, there is not an overall agreement on which category is best for communicating risks. For example, Karelitz & Budescu (2004) find that most people prefer verbal communications. However words can be interpreted differently, and therefore other authors (see for

example Behn & Vaupel (1982), Von Winterfeldt et al. (1986)) recommend numerical probabilities. Price et al. (2007) find supportive evidence for the use of graphical representations of frequency information. Instead of making general recommendations, Duklan & Martin (2002), Visschers et al. (2009) and Spiegelhalter et al. (2011) point out that the context in which the risk communication takes place plays an important role.

Additional literature shows that companies use presentation forms to present information in a self-serving manner, for example to give a more favorable view on the performance (Bettman & Weitz (1983), Beattie & Jones (2000)). Presenting information to nudge people in a certain direction is based on a concept known from psychology, the framing effect. The work of Tversky & Kahneman (1981) has shown that people make different choices depending on whether the risky choice is formulated positively or negatively (e.g. save people's life/let people die). Later, Levin et al. (1998) summarized the framing effects in three categories: *risky choice framing*, *goal framing* and *attribute framing*, where the latter one is of most interest to us. It describes the situation where a certain attribute is framed, e.g. beef is described as 75% lean instead of 25% fat while people rated the beef with the former description higher than the latter one (Levin & Gaeth (1988)).

Since the introduction of the modern portfolio theory by Markowitz (1952) in the finance literature, risk has been considered equivalent to variance. Risk communication is also often about variance, for example the SRRI in the current KIID, a simplified scale of riskiness from one to seven to guide and investor's risk perception, is based solely on the volatility of weekly returns. In the meantime several authors in different fields agree that perceived risk is in-fact multidimensional. Olsen (1997) suggests that investment risk is a function of four attributes, and shows that the expectation of a loss is the most important attribute in determining risk perception. In addition, Lipkus (2007) finds that it includes both the probability and severity of a loss. While researching people's understanding of financial risk associated with retirement in Britain, Vlaev et al. (2009) also find that risk perception is multidimensional, including volatility and economic uncertainty. When evaluating the risk measures used by individual investors, Veld & Veld-Merkoulova (2008) find that most use more than one measure, and that semi-variance (i.e. variance of negative deviations) is the most popular, and several other shortfall measures are also used. Since Weber et al. (2005) and Veld & Veld-Merkoulova (2008) have shown that risk perception significantly influences risk taking behavior, we also test how the presentation forms effects the risk perception and behavior.

Our study shares similarities with Weber et al. (2005), who find that expected returns and perceived risk, not expected volatility, best predicted portfolio decisions. They examine how the presentation format of investment options affect investors' expectations about risk, returns, and volatility and how these expectations influenced investment decisions. However, they did not compare bar charts with the commonly used line charts, they did not evaluate the investors' ability to compare funds in their

analysis, or their expected potential loss. An analysis of the socioeconomic characteristics is also not given in their paper.

The study of Diacon & Hasseldine (2007) is most similar to this paper. They compare return bar with price line charts, and additionally test the effect of time horizons shown in the past performance chart. However, their study only uses a subjective measure for the risk perception and no objective criteria like the accuracy of the return, volatility and loss expectations. They also did not test how well people could compare funds with a given presentation format, nor did they control for the effect of historical performance paths (De Bondt (1993)). Furthermore their study design did not allow the participants to vary the proportion of equities and bonds in the portfolio, rather they had to choose one or the other. They furthermore did not test alternative presentation formats, such as both bar and line chart together, or otherwise.

3 Research Hypotheses

Given the insights from the existing literature, in addition to comparing the results of return bar charts and price line charts, we first test whether we can confirm the results of Weber et al. (2005) and Veld & Veld-Merkoulova (2008). This includes testing whether subjectively perceived risk on a scale of one to seven can in-fact best explain the risk taking behavior. We also test whether additional information like the SRRI (used in the KIID) and a range of gains and losses around the price line we propose will improve the investor's understanding of the return distribution and the loss estimation.

Within the context of the KIID, our primary research hypotheses are:

H1 *Which risk communication leads to the most accurate assessment for a) return, b) volatility and c) loss expectations and d) the highest perceived (past and future) risk?*
H2 *Which risk communication improves the comparability of funds?*
H3 *Which criteria have the most influence on investment decisions? The criteria are: risk & reward, past performance, perceived (past) risk, perceived (future) risk, average return, highest/lowest return, expected loss, last return, past experience and finally theoretically highest/lowest possible historical fund returns.*
H4 *Is there a relationship between the accuracy of an investor's risk assessment and his a) socioeconomic characteristic, b) financial knowledge and c) confidence?*

With regard to the first hypothesis, we would expect that price line charts lead to a more accurate assessment of the return, volatility and loss expectation, since return bar charts can result in framing effects as the cumulative return or the end price are not easily apparent. Furthermore we expect that the price line chart including the range of gains and losses best communicate the potential for losses. As risk perception is subjective, its accuracy cannot be assessed. However, given the regulatory

objective of informing investors so that they do not underestimate the risks, we measure the effectiveness of the risk communication by how high the risk is perceived to be. Since the return bar chart makes negative and positive outcomes more visible (returns can be both positive and negative, while prices are always positive) we expect that the perceived risk is higher for return bar charts.

With the second hypothesis we analyze how well participants can compare the risks of different funds. We question whether that the simple ranking of the Synthetic Risk and Reward Indicator (SRRI) in fact contains meaningful information to investors by supporting them in comparing funds which appear different but have similar return volatilities.

With regard to the third hypothesis we test which of the criteria (return, volatility and loss expectations, the risk perception, etc.) have the most observed and self-reported influence on investment decisions. As the objective of a regulator is to best communicate those criteria which in-turn nudge the investor to not take too much risk; we aim to identify these criteria. Since Weber et al. (2005) and Veld & Veld-Merkoulova (2008) have shown that risk perception significantly influences risk taking behavior, we would also expect to confirm those results. We additionally test whether any significant relationships between socioeconomic characteristics, knowledge or confidence on an investor's risk and return assessments can be observed.

4 Survey design and data sample

The survey is designed to test and challenge the risk communication used in the KIID. To this end, five treatments, each with different presentations of historical performance, are compared. The usefulness of additional information, such as the SRRI as well as a range of gains and losses around the price line are also tested. The risk communication which supports the participants in giving the most *accurate* assessments of the risk and return characteristics of a fund, and in *comparing* risky funds correctly, is considered the most effective. The importance of these criteria in actually determining the investment decisions is also considered.

4.1 Method and subjects

An online survey was conducted in October 2014 with 1000 participants, i.e. 200 per treatment. The participants are a broad sample of the UK population, and were recruited by Research Now. Socioeconomic questions were used to apply a quota sampling procedure for selecting participants to ensure a broad sample for each treatment. The data is cleaned with the use of plausibility checks. For example, the response for the highest expected price in one year must be above the lowest expected price. Likewise participants who gave extreme responses, such as an expected return of 900%, were screened out. After the data cleaning the sample was reduced to 637.

The participants were monetarily incentivized for taking part in the survey. The final payment included a fix and a variable component. The fix payment was required by Research Now and

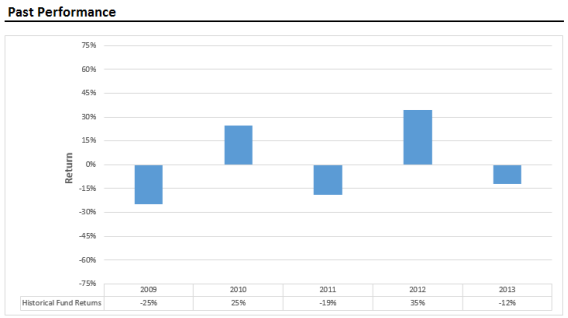
consisted of 1.96 pounds. The variable payment was additionally given and ranged between 0 and 1.6 pounds. For the variable payment one of the three incentivized investment decisions which the participants made during the survey is randomly selected, and a return from its return distribution is randomly drawn. The participants received an average total payout of 2.46 pounds, which is given the average completion time of 10 minutes and 30 seconds, equivalent to an hourly salary of 14.07 pounds. Given the average hourly salary of 13.03 pounds in the UK (Office for National Statistics (2013)), the survey can be considered well incentivized.

4.2 Treatments

The design contains five treatments, each with a different risk communication (see Figure 1). Treatment one is a simple bar chart of return, treatment two is the same with the addition of the SRRI. Treatment three is a simple price line chart. Treatment four is also a price line chart with the additional lines representing the 10% and 90% percentile gain and loss range. Lastly, treatment five is a combined chart with both the return bars and the price line.

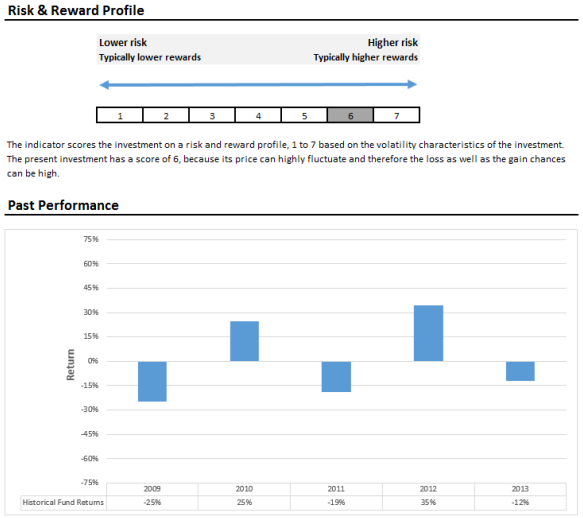
Figure 1: The five treatments

Treatment 1: Return bar chart



Past performance is not a guide to future performance. The value of your investment and income from it may go down as well as up and you may not get back the amount you invested.

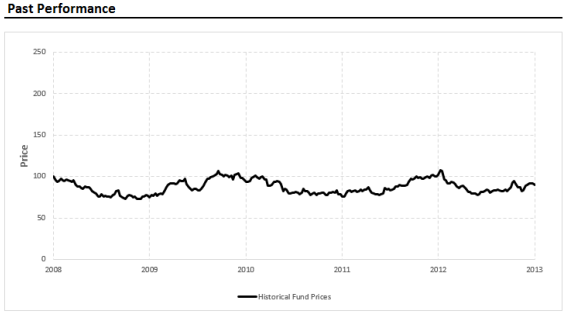
Treatment 2: Return bar chart and SRRI (current KIID)



The indicator scores the investment on a risk and reward profile, 1 to 7 based on the volatility characteristics of the investment. The present investment has a score of 6, because its price can highly fluctuate and therefore the loss as well as the gain chances can be high.

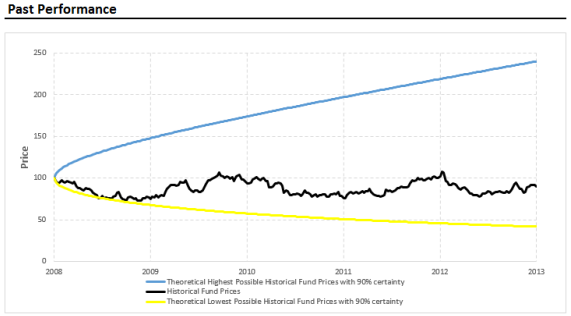
Past performance is not a guide to future performance. The value of your investment and income from it may go down as well as up and you may not get back the amount you invested.

Treatment 3: Price line chart



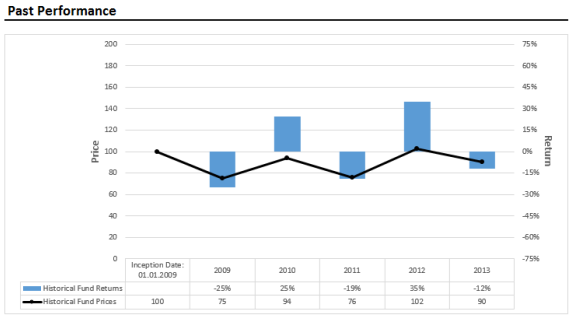
Past performance is not a guide to future performance. The value of your investment and income from it may go down as well as up and you may not get back the amount you invested.

Treatment 4: Price line and range chart



Past performance is not a guide to future performance. The value of your investment and income from it may go down as well as up and you may not get back the amount you invested.

Treatment 5: Return and price chart



Past performance is not a guide to future performance. The value of your investment and income from it may go down as well as up and you may not get back the amount you invested.

While it is common in industry to view bar charts with an annual frequency, this is not the case for price line charts. In this survey we therefore use an annual frequency for bar charts and a weekly frequency for price line charts. In treatment five where the price and return information is combined the frequency is consistent, annual. It is also worth noting that while the graphical bar chart in the KIID uses an annual frequency, the SRRI is based on a weekly return frequency.

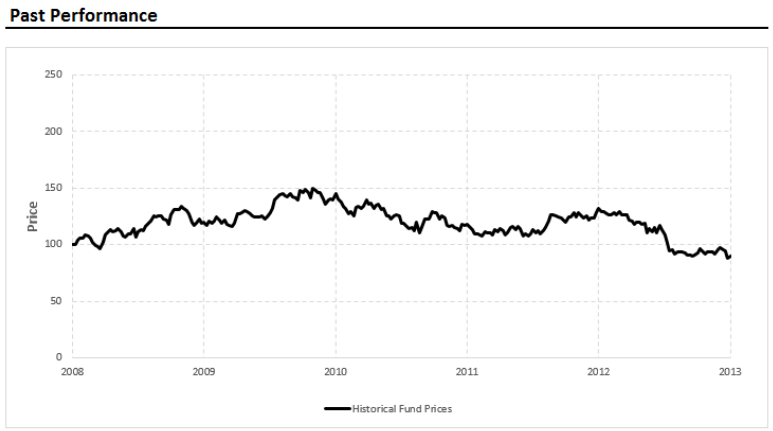
To avoid any further framing effects, all bar charts have the same scale, and all line charts have the same scale. The KIID requires a minimum of five years of historical performance, and so all treatments use this horizon. To replicate the context of the KIID, the treatments with bar charts include a table giving the historical fund returns. The intention of this design is to measure how well the participants can assess the risk and return characteristics of a fund over the next year for each of the communication formats.

4.3 Within the Treatments

Given the results of various authors (Ippolito (1992), De Bondt (1993), Hendricks et al. (1993), Chevalier & Ellison (1997), Sirri & Tufano (1998), Choi et al. (2010), Grosshans & Zeisberger (2015)) that the historical performance path (e.g. whether it is up-down, down-up or straight) can influence the investment behavior, it may also influence the accuracy of the risk and return assessments and risk taking behavior. We have therefore controlled for this effect in the survey design and further used it to test the comparability of funds.

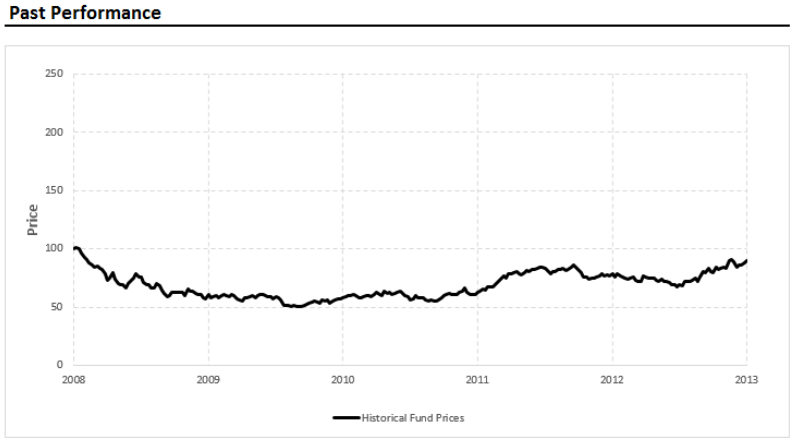
Figure 2: Three funds, each with a different price path, shown within each treatment. Treatment 3 is used here as an example

Path: Up-Down



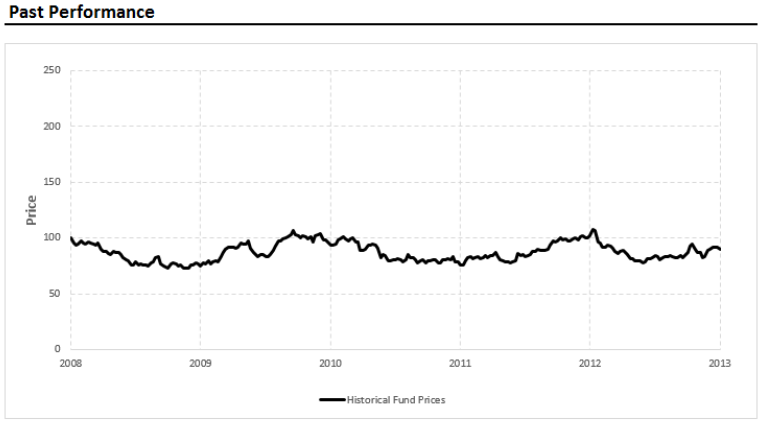
Past performance is not a guide to future performance. The value of your investment and income from it may go down as well as up and you may not get back the amount you invested.

Path: Down-Up



Past performance is not a guide to future performance. The value of your investment and income from it may go down as well as up and you may not get back the amount you invested.

Path: Straight



Past performance is not a guide to future performance. The value of your investment and income from it may go down as well as up and you may not get back the amount you invested.

Within each treatment three graphs are shown, each with a different past performance path: up-down, down-up, and straight (see Figure 2). To the participants each graph is communicated as belonging to a fund. As it will later be shown, the accuracy of the risk assessment is calculated as the average accuracy bias over the three paths (i.e. controlling for the path influence). The development paths are selected from a Monte Carlo simulation with 10'000 simulated paths which are based on a geometric Brownian motion with zero drift and 20% standard deviation. It was carefully checked that all three selected paths have the same risk and return characteristics. A participant who recognizes that the three funds they view come from the same return distribution and simply have different performance paths, is better able to compare risky funds. The treatment which best mitigates the path influence can be considered superior, since it offers the best comparability of funds. We control for order effects by randomizing the order of the three paths.

4.4 Dependent and measured variables

For each of the three fund graphs, the participants answer questions on their risk perception and their risk and return estimates. The participants rate the risk of the past and expected future performance on a scale from one (very little risk) to seven (very high risk). The rating indicates the perceived risk of each fund. They additionally give their estimates of the annual expected median fund return, and the highest/lowest fund return (the 10% and 90% percentile) over a horizon of one year. For each fund a participant views, they make an investment decision. The investment decision is to allocate a fictitious 100'000 pounds (equivalent to 50 pence) between cash with a return of zero, and the fund with an unknown return.

At the end of the survey the participants are directly asked which criteria they considered to be most important in making investment decisions. The ten criteria to choose from are: risk & reward, past performance, perceived (past and future) risk, average return, highest/lowest return, expected loss, last return, past experience and the theoretical highest/lowest possible historical fund returns. Importance is given on a scale from one to seven. The knowledge and confidence of the participants is also asked in this survey (as in Choi et al. (2010)), in order to compare the responses on the return, volatility and loss expectation biases as well as their risk perception.

4.5 Bias measurement

We evaluate the different risk communication formats against the accuracy of the participants' responses. The accuracy is assessed by measuring the participants' assessment bias of the expected return, volatility, and loss. A bias is defined as the deviation from the true value. The true values are given by the defined return distribution used for the geometric Brownian motion (see section 3.3), namely $\mu = 0\%, \sigma = 20\%\ and\ 1\%\ percentile\ (expected\ maximum\ loss) = -47\%$. In other words, we compare participants i's estimated return distribution with the distribution of the past

returns shown. The accuracy of the perceived (past and future) risk is also measured, and as it is subjective the relative score is considered.

In a first step, for each of the three past performance paths *j* (up-down, down-up, straight) investor i's expected return as well as volatility is approximated with the three point approximation proposed by Pearson and Tukey (see Keefer & Bodily (1983)). The approximation is done with the investor i's expected median return ($Y_{ij}^{0.5}$) as well as the expected 10% ($Y_{ij}^{0.1}$) and 90% ($Y_{ij}^{0.9}$) percentile returns. The exact formulas can be found in Appendix I. No calculations are done for the expected absolute lowest possible fund return ($Y_{ij}^{0.01}$) since we directly asked for it.

In a second step the participant i's assessment bias of the expected return, volatility and loss is calculated for each of the three past performance paths *j*. The true value of the presented past return distribution $\mu = 0\%, \sigma = 20\%\ and\ 1\%\ percentile\ (expected\ maxium\ loss) = -47\%$ is subtracted from the investor i's expectation to calculate the bias. For example, if participant i's expected volatility is 12% his assessment bias of the expected volatility would be -8%, as the true volatility is 20%. Therefore the participant would underestimate the volatility.

In a last step we control for the path influence by taking the average assessment bias for the expected return, volatility and loss over the three past performance paths *j*. This last step is also done for the perceived risk. For example, if participant i's assessment bias of the expected return is 6% for the down-up, 2% for the straight and 4% for the up-down past performance path, the average assessment bias is 4%. In other words, the participant overestimates the return by 4%.

4.6 Comparability measurement

How well investors can compare funds is an important criteria in evaluating risk communication. We test which of the treatment groups could best recognize that the three past performance paths (which were communicated as individual funds) have in fact the same return distribution. The degree of comparability is measured by calculating the standard deviation in the assessment biases of the expected return, volatility and loss as well as in the perceived (past and future) risk between the three past performance paths (see Appendix II); the lower the standard deviation in the participant's assessment biases, the better the comparability.

For example, assume that participant i's assessment bias of the expected loss is -2% for the down-up, -4% for the straight and -6% for the up-down past performance path. In this case the standard deviation of the participant's responses would be 2%, and hence he does not fully recognize that the expected loss of the three past performance paths are the same.

4.7 Statistical Tests

Differences between the treatments regarding the accuracy, comparability and path influence are based on their central tendencies, and are statistically tested with the nonparametric Kruskal-Wallis-Test. As a post-hoc test we use the Wilcoxon-rank sum test. As the multiple pairwise comparisons of the post-hoc test causes alpha-inflation, we adjust the significance level with the Bonferroni correction:

$$\alpha_k = \frac{\alpha}{k}$$

$$with\ k = number\ of\ pairwise\ comparison$$

In our case this results in a corrected significance level of $\alpha_5 = 0.01$ which is used instead of the standard level $\alpha = 0.05$ to decide if the differences between the treatments are significant or not.

The criteria which have the most influence on investment decisions is evaluated in two ways. Firstly, the data on the self-reported criteria importance is analyzed. Participants ranked the criteria on an ordinal five-point scale from one 'not important at all' to five 'very important'. Secondly, an OLS regression analysis of the independent variables expected return, volatility and loss, as well as the perceived (past and future) risk, is made on the investment decisions (see Appendix III).

Finally, the relationship between the accuracy of a participant's assessment of the risk and return characteristics, and his own socioeconomic characteristics, financial knowledge and confidence are measured by analyzing the descriptive data and by doing multiple linear regression analyses.[3] By using a dummy variable for the five different treatments we control for any treatment effects (see Appendix IV).

5 Results

The results are organized by our research questions, namely by the accuracy of a participants` assessment of the risk and return characteristics, the comparability of risky funds, and the importance of the decision making criteria. The additional analyses on socioeconomic variables, knowledge and confidence, as well as the influence of performance path are also presented.

5.1 Accuracy of the risk and return assessments

Sizable differences in the assessment biases, i.e. deviation from the true value for return, volatility and loss as well as differences in the perceived (past and future) risk can be found between the five treatments as the Kruskal-Wallis-Test is in all cases significant ($p < 0.001$). Table 1 shows the results of the post-hoc test.

[3] Since some of the dependent variables are non-interval variables (perceived past and future risk) we also did an ordinal logistic regression analysis for those cases – the results remain constant.

As expected, the return and volatility assessment is better for those participants who viewed the price line charts than those who viewed the return bar charts. The participants who viewed the price line charts three and five have an average positive return expectation bias of 4.67%, while those who viewed only the return bar charts have an average bias of 6.77%. The differences are however only significant on the α=0.05 level, but are insignificant on the Bonferroni corrected α= 0.01 level.

The expected volatility bias for the price line treatments three and four is on average 4.47%, and is significantly lower compared to the other treatments ($p<0.001$)[4]. In comparison treatments one, two and five (with a return bar chart) have on average an expected volatility bias of 8.25%. In other words, the participants who viewed the return bar charts underestimated the volatility the most. We find therefore that there is most likely a framing effect resulting from the use of bar charts, as the participants in those treatments have more difficulty in estimating the expected return and volatility.

The participants in treatment four, the price line and range chart, have a positive return expectation bias of 15.43%, and thus extremely overestimated the expected return. This bias is likely due to the unsymmetrical effect of the return on the price. Given equal gain and loss probabilities, after multiple periods the highest price deviates further from the original price than the lowest price. The treatment four chart visually shows this effect. By starting at 100, after five years the upper boundary price is approximately 250 while the lower boundary price is approximately 50.

When asked to estimate the absolute lowest possible fund return (or price, depending on the treatment) in one year, the participants underestimate the loss significantly. While the expected loss in a 99% Value at Risk situation is -47%, the participants in all treatments, except one, estimated it to be between -8.17% and -16.33% (i.e. less negative). The exception is for treatment four (the price line chart with gain and loss range), where the loss expectation is more accurate, at -25.42% ($p < 0.0001$ compared to the other treatments).

Although the price line and range chart (treatment four) resulted in a much improved (in significance and magnitude) assessment of the potential loss compared to all other treatments, it also has some draw-backs. Firstly, the asymmetrical effect of return on price developments leads investors to extremely overestimate the expected return. Secondly, as it will be shown later the fund comparability is worse, and the perceived risk is low compared to the other treatments (except treatment three).

The participants in the return bar chart treatments one and two have a significantly ($p < 0.001$) higher risk perception than those in the price line charts in treatments three and four. The perceived risk is highest, and closest to the risk category as calculated in SRRI in treatment two, the current KIID. The SRRI therefore significantly helps to further improve the risk perception, relative to all other treatments. However in all treatment groups the participants estimate the risk to be lower than as

[4] With exception to treatment three relative to treatment 1, where the p-value is only smaller than 0.1.

calculated in the SRRI; while the risk ranking is six according to the SRRI calculation, the median perceived (past and future) risk ranges from 3.67 to 4.67[5].

Overall the participants underestimate the expected volatility, loss and perceived risk, and overestimate the expected return. The underestimation is especially very large for losses. Although the responses of the participants are not accurate, there are notable differences between the treatment groups, indicating that different risk communication forms can influence the accuracy of investor's risk and return assessments. The current KIID (treatment two) results in the highest perceived risk, while the price line charts result in the most accurate return and volatility assessments. Finally, the price line chart with the gain and loss range (treatment four) best supports the most accurate loss expectation.

5.2 Comparability

Regarding fund comparability the Kruskal-Wallis-Test shows that there are significant differences between the treatments with respect to the three assessment biases (return, volatility and loss), and the perceived (past) risk ($p<0.05$), however not for the perceived (future) risk. Table 2 shows the results of the post-hoc test.

The participants in treatment five with the combined return and price chart have the most consistent return, volatility and loss expectations, which is also significant on a 1% level with respect to treatments 1 and 3. Treatment five is also significantly more consistent than treatment 4 (price line chart with the gain and loss range), except for with regard to the loss expectation. Treatment five is not significantly better than treatment 2 (return bar chart and SRRI).

The return bar chart and SRRI, treatment two, has the second lowest (median) standard deviation in the responses for all criteria (except for perceived future risk), however compared to the other treatments the result is in most cases not significantly better. For perceived past risk, treatment two has the lowest standard deviation which is also significant on the 1% level (except for relative to the price line chart, which is significant on the 5% level).

5.3 Important decision making criteria

The actual influence of the criteria analyzed (e.g. return assessment) on the investment decisions is presented herein. Firstly, with the OLS regression we analyze which of the independent variables (expected return, volatility and loss as well as perceived (past and future) risk) can best explain the incentivized investment decision. The results in Table 3 show that the signs are consistent with what one would expect; a higher expected return leads the participants to take more risk, while a higher expected volatility, perceived risk and larger expected loss lead investors to take less risk. The only statistically significant results on the 1% level are for the perceived (past and future) risk with and

[5] Unlike the return and volatility assessments, the risk perception is limited to a scale from one to seven and therefore the differences in assessment biases is smaller between the treatments.

average adjusted R-squared of 5.3% and 7.6% respectively. This result is consistent with the findings of Weber et al. (2005), and even with a somewhat higher R-squared. But we could not confirm the further result of Weber et al. (2005) that the expected return could significantly explain the risk taking behavior. However, it is one of the four criteria which the participants reported as most important.

Secondly, we analyze the criteria[6] which the participants rated important. As shown in Table 4 the most important criteria are the perceived (past and future) risk, the past performance and the return expectation, as self-reported by the participants. These four criteria are ranked as most important by the participants in all treatments. The theoretical highest/lowest possible historical fund returns is ranked low in importance (3 out of 9) and the SRRI ranked in the middle (5 out of 9). The self-reported results are consistent with the results from the OLS regression. In other words, the participants are most influenced by those criteria they say they are.

5.4 Socioeconomic characteristics, financial knowledge and confidence

These results relate to the remaining research questions, namely whether socioeconomic variables, financial knowledge and confidence influences the accuracy of a participants' assessment of the risk and return characteristics.

The results from the OLS regression show that only for the expected volatility a significant difference for the socioeconomic variables could be found (see Table 5). Older people tend to have a higher miss-estimation of the expected volatility compared to younger people.

The descriptive data in Table 6 further shows that the relationship between knowledge and confidence and the different assessment biases (expected return, volatility, loss and perceived (past and future) risk is not monotonic, i.e. there is no observable linear relation. Unreported ordinal logistic as well as linear regression analyses confirm that there is no significant relationship.

5.5 Verification of the path influence

We observe that the historical performance path (up-down, down-up and straight) does influence the expected return, volatility, loss and the perceived (past and future) risk, and therefore the control for this effect in the survey design is worthwhile. Furthermore it confirms that our design allowed us to analyze the comparability of funds. As an example from Table 7[7], the expected return is overestimated the least (by 3.2%, $p < 0.0001$) for the fund with the up-down path (across all treatments), compared to the funds with the straight (5%) and down-up (10.5%) paths. To no surprise, the participants underestimate the volatility the most (by -9.04%, $p < 0.0004$) for the funds with the straight path. They underestimate the volatility the least (by -5.23%, $p < 0.0001$) for the funds with the down-up path. For

[6] The criteria are the same for all treatments, with the exceptions that the risk & reward profile is only available in treatment two (return bar chart and SRRI), and the theoretical highest/lowest possible historical fund returns are only available in treatment four (price line and range chart).

[7] Note that the Kruskal-Wallis-Test for the difference in path influence was significant for all variables ($p<0.05$)

the funds with the up-down path they also have the highest past risk perception of 4.35 ($p < 0.001$). Finally the future perceived risk is highest (4.51, $p < 0.0001$) for the funds with the up-down path.

In an unreported regression analysis we find that participants who rate past performance and return expectation as important, also have significantly more variance in their investment decision making. This result is intuitive, as the different past performance paths (up-down, down-up and straight) lead those participants to take different decisions.

As we observed a path influence, we can confirm that controlling for this influence when conducting similar research is important. The influence possibly relates to momentum or contrarian beliefs, which however has little relation to assessing the return distribution with as little bias as possible.

6 Discussion

In this paper we attempt to answer the question of how to best communicate risks in the context of the KIID. For the European regulators we believe this is the central question, and not whether or not past performance should be included. To this end, we compare different graphical representations of historical performance and some additional information.

In line with Diacon & Hasseldine (2007) we find that the past performance is an important decision making criteria for investors. By withholding the information from investors it has to be expected that investors take efforts to see the past performance anyway. This would lead to unnecessary searching cost for the investor which would definitely not increase his satisfaction, and he would still have not learned that the past performance is not an indicator for the future performance. Withholding the information could also cause mistrust in the investor, e.g. that the past performance is withheld because something is wrong with it.

However, past performance may be falsely used by investors as an indicator for future performance. Although the regulator has made efforts to warn investors that the past performance is not an indicator for the future performance, the study of Mercer et al. (2010) has shown that such disclaimers are ineffective. We therefore agree with Druckman (2001) that fund information which could lead to framing effects should be explained by an advisor. The advisor should especially explain information which could mislead the investor.

Our data shows that the participants in all treatment groups underestimate the volatility and potential loss, and overestimate the expected return. On average they underestimate the volatility by 6.7%, the potential loss by 32.8% and overestimated the return by 7.7%. Since those factors are part of defining the perceived risk (Veld & Veld-Merkoulova (2008), Vlaev et al. (2009)) people tend to perceive risk

to be lower than it actually is.[8] This conclusion can also be confirmed with our data, which show that the perceived risk is clearly lower than the reference value of the SRRI given by the regulator. We also confirm the finding of Weber et al. (2005) and Veld & Veld-Merkoulova (2008) that the perceived risk is the most important factor for making an investment decision. Given that people's risk perception is relatively low, and that it is an important factor in decision making, it can be concluded that people take more risk than they actually want to. While each risk communication format tested has its advantages and disadvantages, the bar chart combined with the SRRI (as in the current KIID), leads to higher risk perception compared to the other formats, which nudges people to take less risk.

Our results are also consistent with Lipkus & Hollands (1999) who test visual communications of risk, and find that people have difficulty in understanding the magnitude of small probability events. However, the assessments can be improved with communicating risk in better ways. For example, the results of treatment 4 (price line with loss and gain range) show that when the possibility of a loss is directly communicated, people better understand it.

The lack of results for socioeconomic backgrounds, knowledge and confidence on risk and return assessments or risk taking behavior means that no evidence can be found that one risk communication works better for one demographic than for another. In other words, the same risk communication in the KIID would be equally effective for all. This finding supports the use of one common KIID template to increase the comparability of risky funds.

7 Conclusion

We find that the presentation of past performance does have an effect on the accuracy of participant's assessments for return, volatility and loss expectations and on their risk perception. The return bar charts, as used in the current KIID, supported the highest and most accurate risk perception, which is found to be the most important factor for investment decision making. While an accurate assessment of the expected return and volatility, supported best with a price line chart, is desirable, they are of less importance in decision making.

The Synthetic Risk and Reward Indicator (SRRI) supports the comparability of funds, i.e. the participants who viewed it could reasonably well recognize that the three funds have the same return distribution. We found no patterns for socioeconomic variables, knowledge or confidence and therefore there is no evidence that the same risk communication (i.e. same KIID) would not be equally effective for all demographics.

Given the extreme underestimation of the potential for losses in our results, and the well-established effect of loss aversion on investment decisions, we find that improved communication about losses is

[8] This conclusion can also be confirmed in our data which show that the perceived risk is clearly lower than the reference value of the SRRI given by the regulator.

highly recommendable for the KIID. While the price line chart with a gain and loss range performed best with regard to this criteria, it performed rather poorly in the other criteria and thus may cause undesired effects. An alternative written communication format may better communicate loss and would require further testing in the context of the KIID. We suggest that such a communication presents data both in a loss and gain frame[9], and that it uses frequencies (e.g. 1 in 100) in addition to percentages (e.g. 1%), such as in Fagerlin et al. (2007) and Stössel (2015).

There are limitations to our study. We tested only a part of the KIID and compared it to alternatives, but cannot make a statement as to how the KIID as a whole effects investors understanding and perception of risk. For example, we did not include benchmarks, which are also intended to support comparability of funds. We also could not test the effect of the KIID in combination with explanation and advice from an advisor, as would be the case in an advisory process. Furthermore, we used a normal distribution, and so we also suggest further research on the effectiveness of the KIID for funds with non-normal distributions. We leave these aspects for further research.

[9] Information about the overall return distribution additionally allows for accurate communication of unsymmetrical distributions, the importance of which is mentioned by Weber et al. (2005).

References

Beattie, V. & Jones, M. J. (2000), 'Impression management: The case of inter-country financial graphs', *Journal of International Accounting, Auditing and Taxation* **9**(2), 159–183.

Behn, R. D. & Vaupel, J. W. (1982), *Quick Analysis for Busy Decision Makers*, Basic Books New York.

Bettman, J. R. & Weitz, B. A. (1983), 'Attributions in the board room: Causal reasoning in corporate annual reports', *Administrative Science Quarterly* **28**(2), 165–183.

Chevalier, J. & Ellison, G. (1997), 'Risk taking by mutual funds as a response to incentives', *Journal of Political Economy* **105**(6), 1167–1200. http://www.jstor.org/stable/10.1086/516389

Choi, J. J., Laibson, D. & Madrian, B. C. (2010), 'Why does the law of one price fail? an experiment on index mutual funds', *Review of Financial Studies* **23**(4), 1405–1432.

Dalbar (2011), 'Quantitative analysis of investor behavior. helping investors change behavior to capture alpha', *Dalbar Inc. Boston* .

De Bondt, W. P. (1993), 'Betting on trends: Intuitive forecasts of financial risk and return', *International Journal of Forecasting* **9**(3), 355–371.

Diacon, S. & Hasseldine, J. (2007), 'Framing effects and risk perception: The effect of prior performance presentation format on investment fund choice', *Journal of Economic Psychology* **28**(1), 31–52.

Druckman, J. N. (2001), 'Using credible advice to overcome framing effects', *Journal of Law, Economics, and Organization* **17**(1), 62–82.

Duklan, K. & Martin, M. A. (2002), 'Communicating effectively with words, numbers, and pictures: Drawing on experience', *Journal of Actuarial Practice* **70**, 5–61.

European Parliament and Council (2012), Proposal for a regulation of the european parliament and of the council on key information documents for investment products (/* com/2012/0352 final - 2012/0169 (cod) */), Technical report, European Commission.

Fagerlin, A., Ubel, P. A., Smith, D. M. & Zikmund-Fisher, B. J. (2007), 'Making numbers matter: Present and future research in risk communication', *American Journal of Health Behavior* **31**(Supplement 1), 47–56.

Fischhoff, B. (1995), 'Risk perception and communication unplugged: Twenty years of process', *Risk Analysis* **15**(2), 137–145.

Fischhoff, B., Riley, D., Kovacs, D. C. & Small, M. (1998), 'What information belongs in a warning?', *Psychology & Marketing* **15**(7), 663–686.

Grosshans, D. & Zeisberger, S. (2015), 'All's well that ends well? on the importance of how returns are achieved', *Available at SSRN: http://ssrn.com/abstract=2579636* .

Hendricks, D., Patel, J. & Zeckhauser, R. (1993), 'Hot hands in mutual funds: Short-run persistence of relative performance, 1974-1988', *The Journal of Finance* **48**(1), 93–130.

Ippolito, R. A. (1992), 'Consumer reaction to measures of poor quality: Evidence from the mutual fund industry', *Journal of Law and Economics* **35**(1), 45–70.

Karelitz, T. M. & Budescu, D. V. (2004), 'You say 'probable' and i say 'likely': Improving interpersonal communication with verbal probability phrases', *Journal of Experimental Psychology: Applied* **10**(1), 25–41.

Keefer, D. L. & Bodily, S. E. (1983), 'Three-point approximations for continuous random variables', *Management Science* **29**(5), 595–609.

Levin, I. P. & Gaeth, G. J. (1988), 'How consumers are affected by the framing of attribute information before and after consuming the product', *Journal of Consumer Research* **15**(3), 374–378.

Levin, I. P., Schneider, S. L. & Gaeth, G. J. (1998), 'All frames are not created equal: A typology and critical analysis of framing effects', *Organizational Behavior and Human Decision Processes* **76**(2), 149–188.

Lipkus, I. M. (2007), 'Numeric, verbal, and visual formats of conveying health risks: Suggested best practices and future recommendations', *Medical Decision Making* **27**(5), 696–713.

Lipkus, I. M. & Hollands, J. (1999), 'The visual communication of risk', *JNCI Monographs* **1999**(25), 149–163.

Markowitz, H. (1952), 'Portfolio selection', *The Journal of Finance* **7**(1), 77–91.

Mercer, M., Palmiter, A. R. & Taha, A. E. (2010), 'Worthless warnings? testing the effectiveness of disclaimers in mutual fund advertisements', *Journal of Empirical Legal Studies* **7**(3), 429–459.

Office for National Statistics (2013), 'Annual survey of hours and earnings, 2013 provisional results', *Statistical Bulletin* .

Olsen, R. A. (1997), 'Investment risk: The experts' perspective', *Financial Analysts Journal* **53**(2), 62–66.

Price, M., Cameron, R. & Butow, P. (2007), 'Communicating risk information: The influence of graphical display format on quantitative information perception—accuracy, comprehension and preferences', *Patient Education and Counseling* **69**(1), 121–128.

Sirri, E. R. & Tufano, P. (1998), 'Costly search and mutual fund flows', *The Journal of Finance* **53**(5), 1589–1622.

Spiegelhalter, D., Pearson, M. & Short, I. (2011), 'Visualizing uncertainty about the future', *Science* **333**(6048), 1393–1400.

Stössel, R. (2015), 'Time spent on risk profiling is a good investment', *Work in Progress* .

Tversky, A. & Kahneman, D. (1981), 'The framing of decisions and the psychology of choice', *Science* **211**(4481), 453–458.

Veld, C. & Veld-Merkoulova, Y. V. (2008), 'The risk perceptions of individual investors', *Journal of Economic Psychology* **29**(2), 226–252.

Visschers, V. H., Meertens, R. M., Passchier, W. W. & De Vries, N. N. (2009), 'Probability information in risk communication: A review of the research literature', *Risk Analysis* **29**(2), 267–287.

Vlaev, I., Chater 1, N. & Stewart, N. (2009), 'Dimensionality of risk perception: Factors affecting consumer understanding and evaluation of financial risk', *The Journal of Behavioral Finance* **10**(3), 158–181.

Von Winterfeldt, D., Edwards, W. et al. (1986), *Decision analysis and behavioral research*, Vol. 604, Cambridge University Press Cambridge.

Weber, E. U. & Hilton, D. J. (1990), 'Contextual effects in the interpretations of probability words: Perceived base rate and severity of events', *Journal of Experimental Psychology: Human Perception and Performance* **16**(4), 781–789.

Weber, E. U., Siebenmorgen, N. & Weber, M. (2005), 'Communicating asset risk: How name recognition and the format of historic volatility information affect risk perception and investment decisions', *Risk Analysis* **25**(3), 597–609.

Table 1: Return, volatility and loss expectation biases as well as the perceived (past/future) risk

	Descriptive data			p-value (Wilcoxon rank-sum test. HA: less than)				
Return expectation bias	Number of subjects	Mean	Median	Return bar chart	Return bar chart and SRRI (KIID)	Price line chart	Price line and range chart	Return and price chart
Treatment 1: Return bar chart	126	8.47	6.27		0.1673	0.9700	0.0000	0.9666
Treatment 2: Return bar chart and SRRI (KIID)	122	11.2	7.27	0.8332		0.9953	0.0005	0.9983
Treatment 3: Price line chart	139	2.85	4.89	0.0302	0.0047		0.0000	0.2900
Treatment 4: Price line and range chart	136	12.06	15.43	1.0000	0.9995	1.0000		1.0000
Treatment 5. Return and price chart	114	11.11	4.45	0.0335	0.0017	0.7106	0.0000	
Volatility expectation bias				p-value (Wilcoxon rank-sum test. HA: less or greater than)				
Treatment 1: Return bar chart	126	-5.62	-6.49		0.1500	0.0949	0.0000	0.0131
Treatment 2: Return bar chart and SRRI (KIID)	122	-7.41	-8.66	0.1500		0.0007	0.0000	0.3001
Treatment 3: Price line chart	139	-3.83	-5.12	0.0949	0.0007		0.0000	0.0000
Treatment 4: Price line and range chart	136	15.82	3.82	0.0000	0.0000	0.0000		0.0000
Treatment 5. Return and price chart	114	-8.7	-9.61	0.0131	0.3001	0.0000	0.0000	
Loss expectation bias				p-value (Wilcoxon rank-sum test. HA: less than)				
Treatment 1: Return bar chart	126	35.94	34.18		0.0244	0.9951	1.0000	0.0113
Treatment 2: Return bar chart and SRRI (KIID)	122	41.09	38.67	0.9757		1.0000	1.0000	0.3627
Treatment 3: Price line chart	139	28.14	30.67	0.0049	0.0000		0.9998	0.0000
Treatment 4: Price line and range chart	136	16.12	21.58	0.0000	0.0000	0.0002		0.0000
Treatment 5. Return and price chart	114	43.14	38.83	0.9887	0.6380	1.0000	1.0000	
Perceived (past) risk				p-value (Wilcoxon rank-sum test. HA: greater than)				
Treatment 1: Return bar chart	126	4.24	4.33		1.0000	0.0000	0.0007	0.2613
Treatment 2: Return bar chart and SRRI (KIID)	122	4.8	4.67	0.0000		0.0000	0.0000	0.0000
Treatment 3: Price line chart	139	3.48	3.67	1.0000	1.0000		0.9971	1.0000
Treatment 4: Price line and range chart	136	3.82	4	0.9993	1.0000	0.0029		0.9948
Treatment 5. Return and price chart	114	4.17	4.17	0.7393	1.0000	0.0000	0.0052	
Perceived (future) risk				p-value (Wilcoxon rank-sum test. HA: greater than)				
Treatment 1: Return bar chart	126	4.29	4.33		0.9973	0.0000	0.0001	0.0469
Treatment 2: Return bar chart and SRRI (KIID)	122	4.64	4.67	0.0027		0.0000	0.0000	0.0000
Treatment 3: Price line chart	139	3.54	3.67	1.0000	1.0000		0.9955	1.0000
Treatment 4: Price line and range chart	136	3.84	4	0.9999	1.0000	0.0046		0.9830
Treatment 5. Return and price chart	114	4.13	4	0.9533	1.0000	0.0000	0.0171	

Table 2: Standard deviation of the expectation biases and the perceived (past/future) risk

	Descriptive data			p-value (Wilcoxon rank-sum test. HA: less than)				
Return expectation bias (σ)	Number of subjects	Mean	Median	Return bar chart	Return bar chart and SRRI (KIID)	Price line chart	Price line and range chart	Return and price chart
Treatment 1: Return bar chart	126	12.13	9.55		0.8617	0.0679	0.0877	0.9960
Treatment 2: Return bar chart and SRRI (KIID)	122	10.21	7.21	0.1387		0.0018	0.0043	0.9561
Treatment 3: Price line chart	139	12.82	11.77	0.9323	0.9982		0.4619	1.0000
Treatment 4: Price line and range chart	136	15.49	10.78	0.9126	0.9957	0.5387		1.0000
Treatment 5. Return and price chart	114	9.33	6.42	0.0041	0.0441	0.0000	0.0000	
Volatility expectation bias (σ)				p-value (Wilcoxon rank-sum test. HA: less than)				
Treatment 1: Return bar chart	126	6.47	5.31		0.7776	0.5368	0.0215	0.9966
Treatment 2: Return bar chart and SRRI (KIID)	122	6.17	4.62	0.2229		0.2202	0.0040	0.9789
Treatment 3: Price line chart	139	6.04	5.06	0.4639	0.7803		0.0127	0.9986
Treatment 4: Price line and range chart	136	10.18	7.27	0.9786	0.9961	0.9873		1.0000
Treatment 5. Return and price chart	114	4.95	3.13	0.0035	0.0212	0.0014	0.0000	
Loss expectation bias (σ)				p-value (Wilcoxon rank-sum test. HA: less than)				
Treatment 1: Return bar chart	126	12	10		0.8080	0.0103	0.5247	0.9913
Treatment 2: Return bar chart and SRRI (KIID)	122	11.62	8.66	0.1924		0.0009	0.2112	0.9241
Treatment 3: Price line chart	139	14.14	13.06	0.9897	0.9991		0.9744	1.0000
Treatment 4: Price line and range chart	136	13.92	8.84	0.4759	0.7893	0.0257		0.9777
Treatment 5. Return and price chart	114	9.95	7.64	0.0087	0.0762	0.0000	0.0224	
Perceived (past) risk (σ)				p-value (Wilcoxon rank-sum test. HA: less than)				
Treatment 1: Return bar chart	126	1.14	1		0.9981	0.7848	0.5738	0.2366
Treatment 2: Return bar chart and SRRI (KIID)	122	0.91	1	0.0020		0.0251	0.0066	0.0003
Treatment 3: Price line chart	139	1.11	1	0.2157	0.9749		0.3636	0.0962
Treatment 4: Price line and range chart	136	1.11	1.15	0.4268	0.9934	0.6370		0.2405
Treatment 5. Return and price chart	114	1.18	1.15	0.7640	0.9997	0.9040	0.7601	
Perceived (future) risk (σ)				p-value (Wilcoxon rank-sum test. HA: less than)				
Treatment 1: Return bar chart	126	1.09	1		0.8349	0.5156	0.9884	0.3965
Treatment 2: Return bar chart and SRRI (KIID)	122	1.03	1	0.1655		0.1594	0.8445	0.1000
Treatment 3: Price line chart	139	1.12	1	0.4850	0.8410		0.9845	0.3708
Treatment 4: Price line and range chart	136	0.94	1	0.0117	0.1559	0.0155		0.0047
Treatment 5. Return and price chart	114	1.12	1	0.6043	0.9003	0.6299	0.9953	

Table 3: Percentage of asset allocation explained by the return, volatility and loss expectation as well as the perceived (past/future) risk

Independent variables	Sign	Average adjusted R-squared	Significance level
Return expectation	+	2.8% [-0.9%-11.6%]	-
Volatility expectation	-	0.7% [-0.8%-7.4%]	-
Loss expectation	+	1.6% [-0.9%-7.3%]	-
Perceived (past) risk	-	5.3% [-0.9%-13.3%]	**
Perceived (future) risk	-	7.6% [0.7%-23.5%]	**

Table 4: Self-reported importance of decision making criteria

	Treatment 1: Return bar chart	**Treatment 2:** Return bar chart and SRRI (KIID)	**Treatment 3:** Price line chart	**Treatment 4:** Price line and range chart	**Treatment 5:** Return and price chart	**Average Ranking (1=lowest; 9=highest):**
Return expectation	6	6	6	7	7	6.4
Volatility expectation	2	1	1	2	2	1.6
Loss expectation	1	4	1	4	3	2.6
Perceived (past) risk	7	7	7	5	5	6.2
Perceived (future) risk	8	9	8	9	8	8.4
Past performance	5	8	5	8	6	6.4
Risk & Reward		5				5
Theoretical highest/lowest possible historical fund returns				3		3
End value of the fund	3	2	3	6	4	3.6
Past experience with funds	4	3	4	1	1	2.6

Table 5: The influence of socioeconomic characteristics on the return, volatility and loss expectation biases as well as the perceived (past/future) risk

Independent variables	Return expectation bias	Volatility expectation bias	Loss expectation bias	Perceived (past) risk	Perceived (future) risk
Return bar chart and SRRI (KIID)	2.3016	-1.0982	3.9782	0.53097***	0.33986**
	(3.5267)	(2.0561)	(3.4315)	(0.12751)	(0.12329)
Price line chart	-5.9787	2.8809	-9.3895**	-0.78599***	-0.7755***
	(3.4381)	(2.0044)	(3.3453)	(0.12431)	(0.12019)
Price line chart and range chart	3.2685	22.2177***	-21.0017***	-0.44129***	-0.45644***
	(3.4371)	(2.0038)	(3.3443)	(0.12427)	(0.12016)
Return and price chart	2.2535	-2.8293	6.5709	-0.0926	-0.1716
	(3.6001)	(2.0989)	(3.5029)	(0.13016)	(0.12586)
25-34	-4.0028	-6.2118*	3.1754	0.1119	0.1123
	(4.2996)	(2.5067)	(4.1836)	(0.15546)	(0.15031)
35-44	-1.9717	-9.8984***	9.8814*	0.0496	0.0652
	(4.2346)	(2.4688)	(4.1203)	(0.1531)	(0.14804)
45-54	-2.4520	-9.8973***	8.8061*	0.1640	0.1931
	(4.0565)	(2.3650)	(3.947)	(0.14666)	(0.14181)
55-64	-0.6856	-9.0403***	9.7135*	0.1369	0.1680
	(4.1806)	(2.4373)	(4.0678)	(0.15115)	(0.14615)
> 65	-4.7812	-11.2517***	8.5593*	0.1849	0.1951
	(3.9811)	(2.3210)	(3.8737)	(0.14394)	(0.13918)
Female	-0.2436	3.3261*	-4.8835*	0.0672	0.0820
	(2.209)	(1.2878)	(2.1494)	(0.07987)	(0.07722)
Bachelor	-0.5936	3.7591*	-4.5727	-0.1399	-0.0509
	(2.6442)	(1.5416)	(2.5729)	(0.0956)	(0.09244)
Master	-8.4230*	1.9998	-10.3082**	-0.2272	-0.1399
	(3.7765)	(2.2018)	(3.6747)	(0.13654)	(0.13203)
PhD	8.1455	0.8209	8.0057	0.0560	-0.0229
	(6.6487)	(3.8763)	(6.4693)	(0.24039)	(0.23244)
Another education	0.3654	3.8625	-5.5507	0.0604	-0.0267
	(3.4561)	(2.0150)	(3.3629)	(0.12496)	(0.12082)
No education	-2.1428	-1.5751	-0.1693	0.1572	0.1513
	(16.2016)	(9.4457)	(15.7645)	(0.58578)	(0.5664)
Constant	12.2201**	-1.4589	35.1621***	4.16407***	4.15915***
	(4.1201)	(2.402)	(4.0089)	(0.14896)	(0.14404)
Adjusted R squared	0.0080	0.2696	0.1360	0.1580	0.1241

Table 6: Expectation biases and perceived (past/future) risk by participant's confidence and knowledge

	Answer	Proportion (%) of answers	Variance in risk taking	Return expectation bias	Volatility expectation bias	Loss expectation bias	Perceived (past) risk	Perceived (future) risk
How likely is it that you would change your investment decisions if you consulted a professional investment advisor (1=not likely; 7=very likely)?	1	8.16	13003.69	2.73	-6.30	31.36	4.01	4.13
	2	5.34	14101.88	11.98	-2.05	35.94	3.62	3.78
	3	9.42	13258.18	5.61	-1.47	29.04	4.13	4.12
	4	26.53	11557.24	7.38	-1.10	30.35	4.00	3.96
	5	27.32	13814.94	11.48	0.30	32.53	4.06	4.02
	6	13.34	13255.81	8.90	-5.04	35.82	4.33	4.35
	7	9.89	16385.93	13.50	1.00	33.87	4.24	4.16
How confident are you that the investment decisions you made are the right ones for you (1=Not at all confident; 7=Very confident)?	1	5.18	10985.06	14.48	-0.98	35.70	4.44	4.30
	2	11.46	15336.98	8.83	0.36	29.36	4.05	3.98
	3	18.84	12327.03	10.74	-0.49	33.62	4.13	4.11
	4	30.77	10942.04	5.50	-3.49	31.00	4.02	3.96
	5	19.94	14352.09	8.82	-1.95	32.45	4.10	4.14
	6	10.20	16570.22	10.58	1.11	30.70	4.06	4.16
	7	3.61	20061.17	19.10	-2.59	44.14	3.83	4.00
How knowledgeable as an investor do you consider yourself to be (1=Not at all knowledgeable; 7=very knowledgeable)?	1	14.44	15561.79	13.74	1.34	33.89	4.37	4.29
	2	11.15	15888.80	6.75	-0.15	27.98	4.06	4.08
	3	19.00	13640.04	9.37	-0.45	31.84	4.06	4.00
	4	25.59	10937.81	9.14	-4.01	35.04	3.98	4.03
	5	19.00	13688.38	6.15	-2.92	31.14	4.03	4.07
	6	8.95	11881.69	7.53	-0.63	29.55	4.04	3.96
	7	1.88	11710.72	16.33	0.07	36.94	4.14	3.92

Table 7: Path influence on the expectation biases and the perceived (past/future) risk

	Descriptive data			p-value (Wilcoxon signed-rank test. HA: less than)		
Return expectation bias	Number of subjects	Mean	Median	Up-down	Straight	Down-up
Fund with up-down path	637	5.97	3.2		0.0000	0.0000
Fund with straight path	637	8.11	5	1.0000		0.0000
Fund with down-up path	637	12.93	10.5	1.0000	1.0000	
Volatility expectation bias				p-value (Wilcoxon signed-rank test. HA: greater than)		
Fund with up-down path	637	-2.05	-7.27		0.0004	1.0000
Fund with straight path	637	-3.17	-9.04	0.9996		1.0000
Fund with down-up path	637	0.58	-5.23	0.0000	0.0000	
Loss expectation bias				p-value (Wilcoxon signed-rank test. HA: less than)		
Fund with up-down path	637	18.01	22.89		0.0001	0.0336
Fund with straight path	637	20.37	22.89	0.9999		0.8974
Fund with down-up path	637	19.47	24	0.9665	0.1026	
Perceived (past) risk				p-value (Wilcoxon signed-rank test. HA: greater than)		
Fund with up-down path	637	4.1	4		0.0000	0.9993
Fund with straight path	637	3.78	4	1.0000		1.0000
Fund with down-up path	637	4.35	4	0.0007	0.0000	
Perceived (future) risk				p-value (Wilcoxon signed-rank test. HA: greater than)		
Fund with up-down path	637	4.51	5		0.0000	0.0000
Fund with straight path	637	3.75	4	1.0000		0.9984
Fund with down-up path	637	3.94	4	1.0000	0.0016	

Appendix

Appendix I. Calculation of the bias measurement

A three point approximation proposed by Pearson and Tukey (see Keefer & Bodily (1983)) is made in order to estimate the participant i's return and volatility expectations. For each of the three past performance paths j (down-up, straight, up-down) the approximation is made with the participant i's expected median[10] return ($Y_{ij}^{0.5}$) as well as the expected 10% ($Y_{ij}^{0.1}$) and 90% ($Y_{ij}^{0.9}$) percentile returns.

$$Return(expect)_{ij} = mean_{ij} = 0.3 \times Y_{ij}^{0.1} + 0.4 \times Y_{ij}^{0.5} + 0.3 \times Y_{ij}^{0.9}$$

$$Vol(expect)_{ij} = \sqrt{\left(0.3 \times \left(Y_{ij}^{0.1}/100\right)^2 + 0.4 \times \left(Y_{ij}^{0.5}/100\right)^2 + 0.3 \times \left(Y_{ij}^{0.9}/100\right)^2\right) - \left(mean_{ij}\right)^2}$$

Participant i's assessment bias of the expected return, volatility and loss is then calculated by subtracting the true values $\mu = 0\%, \sigma = 20\%$ and 1^{th} $perc. = -47\%$ from their expectations:

$$Mean(bias)_{ij} = mean_{ij} - 0\ (\mu_{correct})$$

$$Vol(bias)_{ij} = Vol(expect)_{ij} - 20\ (\sigma_{correct})$$

$$Loss(bias)_{ij} = Y_{ij}^{0.01} + 47\ (1^{th}\ percentile_{correct})$$

[10] Average equals the median given the normal return distribution of a geometric Brownian motion. As a robustness check we also used the expected average return directly as the $mean_{ij}$. All the results stay constant. However using the approximation also for the $mean_{ij}$ has the advantage that 1) the $mean_{ij}$ expectation is adjusted by the skewness expectation ($Y_{ij}^{0.1}, Y_{ij}^{0.9}$) and 2) we get no NAN values for the $Vol(expect)_{ij}$ calculation.

Taking the average assessment bias over the three past performance paths j allows one to control for the path influence:

$$\overline{Mean}(bias)_i = \frac{1}{3}\sum_{j=1}^{3} Mean(bias)_{ij}$$

$$\overline{Vol}(bias)_i = \frac{1}{3}\sum_{j=1}^{3} Vol(bias)_{ij}$$

$$\overline{Loss}(bias)_i = \frac{1}{3}\sum_{j=1}^{3} Loss(bias)_{ij}$$

The average perceived risk is directly calculated as:

$$\overline{PerceivedRisk}_i = \frac{1}{3}\sum_{j=1}^{3} PerceivedRisk_{ij}$$

Appendix II. Calculation of the degree of comparability

The degree of comparability is measured by calculating the standard deviation in the assessment biases of the expected return, volatility and loss as well as in the perceived (past and future) risk between the three past performance paths:

$$\sigma(mean(bias))_i = \sqrt{\sigma^2(Mean(bias)_{i1}, Mean(bias)_{i2}, Mean(bias)_{i3})}$$

$$\sigma(Vol(bias))_i = \sqrt{\sigma^2(Vol(bias)_{i1}, Vol(bias)_{i2}, Vol(bias)_{i3})}$$

$$\sigma(Loss(bias))_i = \sqrt{\sigma^2(Loss(bias)_{i1}, Loss(bias)_{i2}, Loss(bias)_{i3})}$$

$$\sigma(PerceivedRisk)_i = \sqrt{\sigma^2(PerceivedRisk_{i1}, PerceivedRisk_{i2}, PerceivedRisk_{i3})}$$

Appendix III. Measure the influence of the criteria on investment decisions

OLS regression analyses of the independent variables expected return, volatility and loss, as well as the perceived (past and future) risk, is made on the investment decisions in order to measure which of the criteria have the most influence on investment decisions:

$$Decision_{ij} \sim \alpha + \beta X_{ij} + e_{ij}$$

$$where\ X_{ij} = perceivedRisk_{ij}, return(expect)_{ij}, vola(expect)_{ij}, loss(expect)_{ij}$$

Appendix IV. Measure the relationship between socioeconomic characteristics and the assessment biases

The relationship between the accuracy of a participant's assessment of the risk and return characteristics, and his own socioeconomic characteristics, financial knowledge and confidence are measured by doing multiple linear regression analyses.[11] By using a dummy variable for the five different treatments k we control for any treatment effects:

$$Y_i \sim \alpha + \beta_1 age_i + \beta_2 gender_i + \beta_3 education_i + \beta_4 Treatment_{i,k=1,\dots 5} + e_i$$

$$Y_i \sim \alpha + \beta_1 confidence_i + \beta_2 Treatment_{i,k=1,\dots,5} + e_i$$

$$Y_i \sim \alpha + \beta_1 financialKnowledge_i + \beta_2 Treatment_{i,k=1,\dots,5} + e_i$$

$$where\ Y_i = \overline{Mean}(bias)_i, \overline{Vol}(bias)_i, \overline{Loss}(bias)_i, \overline{PerceivedRisk}_i$$

[11] Since some of the dependent variables are non-interval variables (perceived past and future risk) we also conducted an ordinal logistic regression analysis for those cases – the results remain constant.